一种基于多尺度带通滤波的洁化算法与GPU实现*

梅盈[1,4]，邓辉[3]，张晓丽[2,1]，王锋[1,3,4]，陈腾达[2]
(1. 中国科学院云南天文台，云南昆明 650011；2.昆明理工大学，云南昆明 650500；3.广州大学天体物理中心，广东广州 510006；4.中国科学院大学，北京 100049)

摘要：中国新一代频电频谱日像仪-明安图射电频谱日像仪（Mingantu Spectral Radioheliograph, MUSER）以高时间、高空间、高频率分辨率工作在0.4GHz~15GHz，为太阳爆发活动初始能量释放区的物理过程、太阳电子加速等研究开辟了新的窗口。高性能高质量太阳成像算法是MUSER数据处理流水线至关重要的研究内容。参考法国墨东天文台太阳干涉阵的数据处理方法，系统讨论分析了多尺度（Multi-Scale）CLEAN算法，给出了适用于MUSER的多尺度CLEAN算法参数，并重点讨论了算法的GPU并行。实验结果表明，改进的多尺度CLEAN在算法效率上比基于GPU实现的Högbom CLEAN提高了近3倍，有效提高了整个MUSER数据处理流水线的性能。

关键词：多尺度洁化；射电成像；图形处理器
中图分类号：TP274 文献标识码：A 文章编号：1672-7673（2018）03

射电观测是研究太阳剧烈活动的重要探测手段。建于中国明安图的螺旋天线阵列-明安图射电频谱日像仪（Mingantu Spectral Radioheliograph，MUSER）在400MHz~15GHz的584个频点上对太阳进行高分辨率成像，其观测将填补日本野边山日像仪（2个频点，17GHz, 34GHz）、法国南茜日像仪（5个频点，150~450MHz）、俄罗斯伊尔库茨克射电日像仪（5.7GHz）等在厘米分米波段的空白区域，对太阳活动及对人类影响的研究起到重要作用[1]。

明安图射电频谱日像仪分为低频阵（MUSER-I, 400MHz~2GHz）和高频阵（MUSER-II, 2GHz~15GHz），其高时间、高空间和高频率分辨率给数据处理系统提出了较大的挑战。高频阵和低频阵的数据接收机每3ms分别接收一次包含16通道的观测数据。在后续的数据处理过程中，如何高效地获取高质量的太阳图像成为数据处理系统的关键问题。

射电干涉阵成图的主要过程包含加权（Weighting）、网格化（Gridding）、傅里叶变换，退卷积等主要步骤，其中退卷积是对脏图进行洁化的过程，是成像过程中最耗时的部分。对日像仪成像来说，退卷积过程也是整个管线设计中最令人关注的问题。本文在细致调研相关工作的基础上，进一步对射电干涉阵成像中的多尺度退卷积方法进行了讨论，给出了相应的实现方法。

1. 相关研究工作

射电干涉成像的基本原理是天空实际亮度是可见度函数的傅里叶变换。由于在实际观测过程中UV覆盖的不完全性，采样函数是离散的。自1974年Högbom提出CLEAN算法的思想后，CLEAN算法得到了极大的发展并在射电望远镜的图像重建算法中得到了广泛的应用。经典Högbom CLEAN算法的关键是为天空实际亮度提供一个模型（），迭代寻找亮度的峰值位置，并如（1）式进行叠加，其中参数为迭代过程中的峰值亮度，为峰值亮度的位置，由残留图像的峰值亮度确定[2-3]。（2）式中的为脏图亮度，为脏束，为公式（1）计算的出的叠加后的亮度值，为残图的亮度，标记迭代次数。

，（1）

.（2）

Högbom CLEAN算法的一些变种算法，如Clark, Cotton-Schwab算法、最大熵算法

*基金项目：国家重点研究发展计划（973计划）(2016YFE0100300)；国家自然科学基金(U1531132,U1631129,U1231205)；国家自然科学基金(11403009, 11463003，1177301)资助.
收稿日期:2017-11-02;修订日期:2017-12-25
作者简介：梅盈，女，博士.研究方向：科学数据处理.Email: meiying@ynao.ac.cn
通讯作者：张晓丽，女，讲师.研究方向：大数据处理及天文信息技术.Email: zhangxiaoli@astrolab.cn

MEM（the Maximum Entropy Method，MEM）随后被提出。Steer Clean 算法被用于日本野边山日像仪的成像，其洁化效率是 Högbom 算法的近 10
[4]。然而，基本的 Högbom CLEAN 算法由于在处理点源和展源上表现的优势，一直被广泛使用。

但随着新一代射电干涉仪的不断发展，在确保基本 CLEAN 算法的简单性和信噪比的前提下，多尺度成为提高效率的方法之一。Multi-Resolution CLEAN, Multi-Scale Maximum Entropy, Wavelets CLEAN 等多尺度算法有效提高了洁化效率，但存在尺度大小固定和计算代价高等问题。

Cornwell 于 2008 年提出的 Multi-Scale CLEAN[3]算法被认为算法直接且具有稳定性和收敛性好等特点。Multi-Scale CLEAN 为减少算法的迭代次数，提高算法的稳定性，采用多个不同尺度大小的卷积核与脏图进行卷积，得到多个卷积图像，在多个图像中寻找全局亮度峰值。Multi-Scale CLEAN 算法的基本流程如下：

（1） 选取不同的尺度大小，与脏图进行卷积，得到卷积后的系列图像；
（2） 寻找全局亮度最大点，记录亮度最大点的位置和所在的尺度大小；
（3） 计算步骤（2）中的尺度与脏束的卷积，再乘以增益值，存储计算结果及尺度大小；
（4） 从（1）中所有的卷积图像中减去（3）的计算结果；
（5） 重复以上 4 个步骤直至到达循环阈值或达到预设循环次数；
（6） 将（3）中存储的结果与洁束卷积并加上参与图得到最终的洁图。

为了控制算法的迭代次数，提高图片质量，Multi-Scale CLEAN 算法需要指定不同尺度的卷积图像不同的权值，每次迭代循环结束后，残图需要乘以所在尺度的权值来增加或减少亮度幅值。经过系列实验，（3）式中权值被认为是较合适的参数值[3]。

，（3）

其中，为权值；a 为当前尺度大小；为最大尺度大小；为权重参数。

为了减小迭代算法对增益的依赖，Cornwell 选择抛物线函数作为 Multi-Scale CLEAN 的卷积核函数如下：

（4）

其中，为球面波函数；（x, y 为像素点的位置坐标）；为尺度大小。当式中时，$m(r, a)$对应 Högbom CLEAN 中的 Dirac 。

由于高斯卷积核与（4）式给出的卷积核函数仅在较高动态范围下存在较小的差异，一般来说，Cornwell 采用不同尺度的高斯卷积核。以洁束大小为基准，卷积核大小选取常用的等比数列，即依次取洁束大小的 1、2、4、8 倍。

为进一步提高处理速度，近几年利用图形处理器技术加快退卷积处理成为主要的方法，如基于图形处理器的网格化算法使得 w-projection 的效率较**中央处理器提高近百倍**[1][5]~~，项目组在前期基于图形处理器的 Högbom CLEAN 算法较中央处理器下的 Högbom 效率提高了十倍以上。但从 Cornwell 提出的多尺度 CLEAN 算法来看，其并行实现存在较大的难度。因为搜索最大值过程虽然可以并行，但需要利用找出的最大值要同时在所有尺度的图像中退卷积，这意味着一个尺度的数据处理必须与其它尺度的数据处理相交叉，影响了并行的实现。~~

2. ~~基于频域的多尺度洁化~~

2. 1 ~~基本思想~~

~~参考 Cornwell 洁化方法，要加快洁化的效率，关键是能够同时进行洁化。实际上，由于 HögbomClean 的基本原理，洁化迭代的过程始终是串行的，要并行的唯一可能性，是同时能够独立地进行多个洁化迭代。~~

~~法国默东天文台太阳干涉阵与 MUSER 一样，也是一个太阳专用的射电干涉阵。其数据处理采用了基于频率带通滤波的多尺度洁化方法[56]。该方法基于两个简化：（1）没有过~~

1 https://arxiv.org/abs/1403.4209

多的展源存在；（2）空间尺度中只有相对较窄的谱。对于太阳观测来说，完全可以满足这两个简化需要。该方法与 Cornwell 多尺度洁化方法有一定的相似性，但更易于并行实现。

参考法国默东天文台的洁化思想，结合明安图射电频谱日像仪的天线分布情况和最长基线情况，最终在本文实现的算法基本原理如下：

（1）在频域 uv 平面上，构建一系列连续的带通滤波器 $F_k(u, v)$，k 为滤波器编号，滤波器的半径以 2 的倍数增长。不同的滤波器对应着不同的空间尺度；与法国不同，在实现时每次将计算最长基线和相应的 uv 分布，反过来计算相应的滤波器带通频率。

（2）通过滤波器，将原来的稀疏 UV 分布 D(u,v)进一步拆分为若干个不同尺度的独立分布，即：；

（3）对于一个不同尺度的 uv 分布，通过逆傅里叶变换（FT）得到脏图：，然后进行常规的洁化处理；

（4）最后的洁图等于各尺度洁图的累加。

2.2 尺度的考虑

从上述的基本实现原理可见，合适的带通滤器器数量是整个算法的关键。在法国默东太阳干涉阵数据处理中，已经开展了前期的研究，共采用了 5 个滤波器。在明安图射电频谱日像仪数据处理中，也借鉴了相同的方法，但进一步做了改进：即根据太阳方位与高度位置，计算实时的 UV 分布，然后根据最远的 UV 点位置，动态计算每个滤波器覆盖的频率。在 MUSER 成图过程中，滤波器一共 5 个，F_0 到 F_4，不同尺度的 UV 滤波设计见图 1。横坐标为频率范围，单位为(动态计算的最长基线除以 60，60 为自定义数据)，由于在日像仪中 UV 覆盖对称，UV 的频率范围相同，在图中用一条曲线表示，对应的 UV 平面如图 2。

根据 5 个不同尺度大小，将脏图分成不同的频率段，随着 *k* 值增大，尺度依次减小，直到覆盖所有的 UV 值。当 *k*=0 时，滤波器为高斯滤波。其尺度为最长 UV 的 20%，从成像来看，这也是最大的尺度。*k*=1、2、3、4 时，滤波器在频率域呈现一个环状，实现一个带通滤波。无论分成多少级，但本质上要确保这些滤波器的和所合成的滤波器在频率上是连续的，覆盖全部 UV 点，如图 1 和图 2 中的最后一张图。在使用高斯滤波的情况下，最终合成的滤波器在边缘逐渐减小的滤波性质，正好减小了由于较大的 UV 处的稀疏导致的成图不可靠信息。

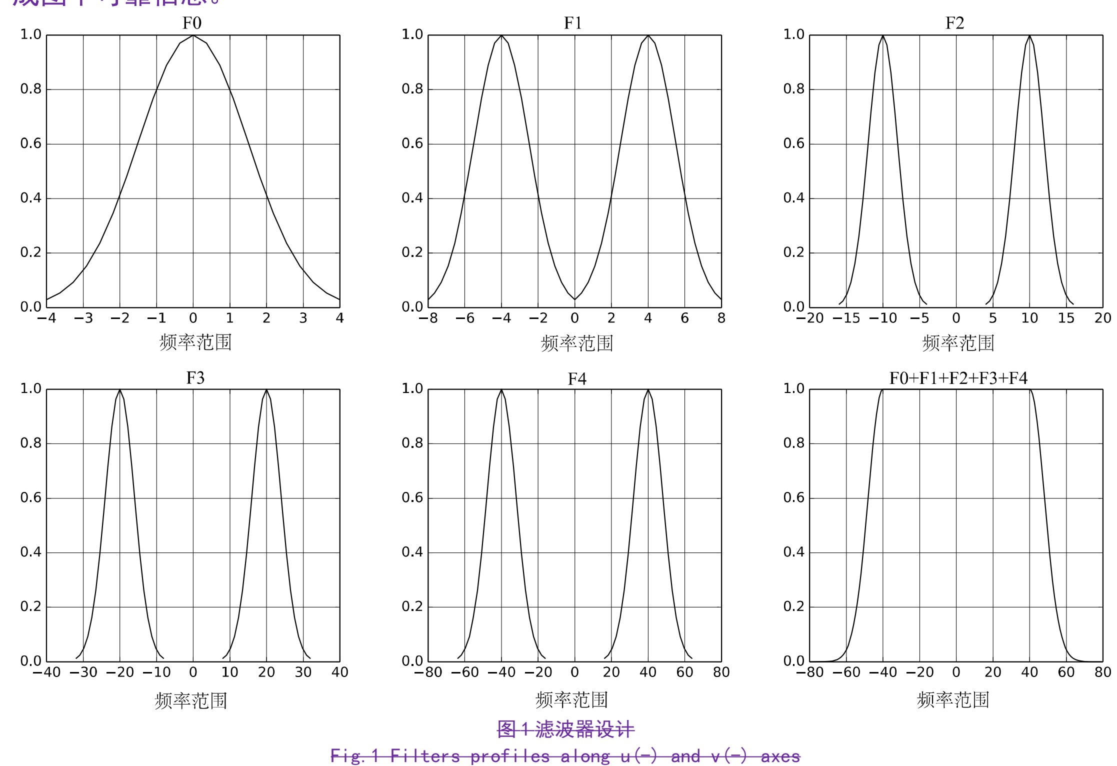


图 1 滤波器设计

Fig. 1 Filters profiles along u(-) and v(-) axes

F0 F1 F2

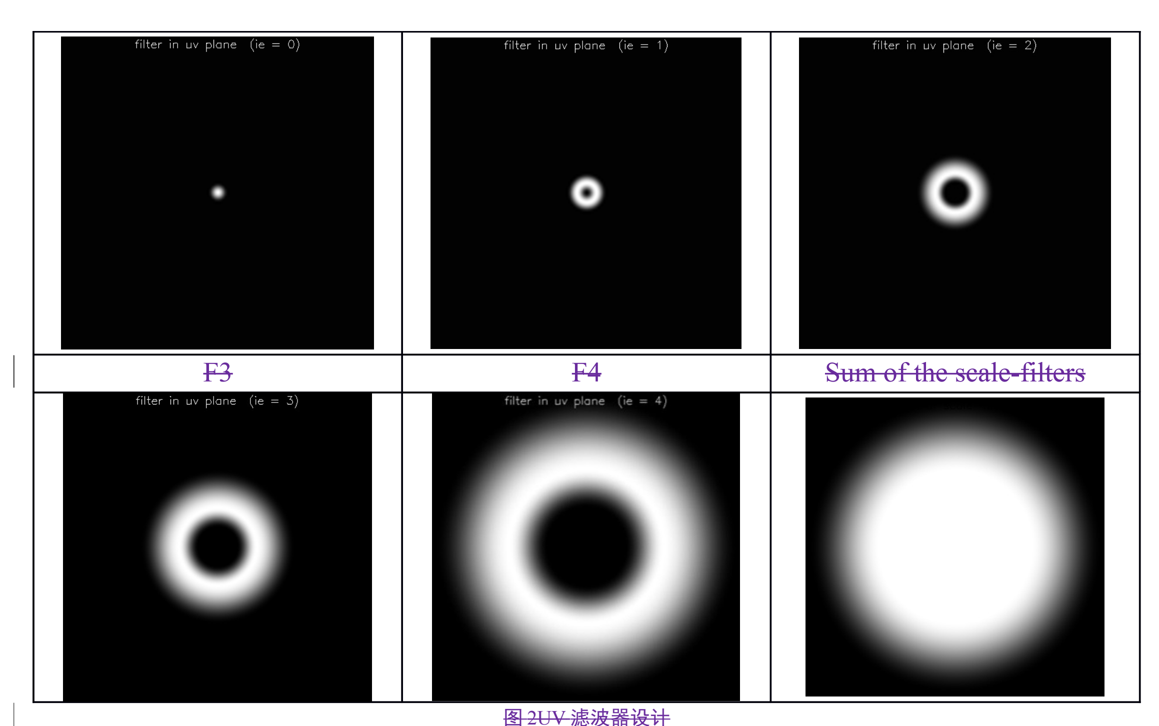


图 2UV 滤波器设计

Fig. 2Filters in the uv-plane according to Figure 2.

3. 算法的图形处理器实现及效率

3.1 图形处理器实现

明安图射电频谱日像仪的成图的图形处理器实现采用 GPU-CUDA 架构，采用 Python 编程语言，PyCUDA，Scikit-CUDA 作为 CUDA 并行编程接口库。实现了数据处理的单机命令行模式及分布式运行模式[67]，其中数据预处理（数据格式转换[78]，相位校准，异常数据标记等）在中央处理器环境下运行，成图及洁化在图形处理器环境下运行，包含 8 台服务器的高性能计算机群（CPUs+GPUs）已在观测站搭建好。成像流程图如图 3，其中 MUSER Multi-Scale CLEAN 算法描述如表 1 。以低频阵观测数据 “2015-11-01 04:08:49.354161240 (UTC), 1.7125 GHz, 右旋” 为例，最终洁化图像如图 4（a）。在实验过程中，CASA 用于验证数据处理各个阶段的正确性。将数据处理系统生成的 UVFITS 文件通过 CASA 转换成 MS 文件并使用 multi-scale 成图可得到基本一致的成图结果。为说明当前处理的正确性，文中给出了日本野边山天文台已经公布的观测处理结果图 4（b）作为对比。由于当前的数据处理过程中没有进行 P 角改正，图像中存在一些误差，在后期的工作中将进行改正。

表 1MUSER 多尺度洁化

Table 1Multi-Scale CLEAN for MSUER

初始化

1. **脏图（dirty image）:，初始情况下两者等价**
2. **计算 uv 覆盖值的范围，选取尺度大小，设计滤波器**
3. **根据原始 uv 和滤波器得到不同频率范围的脏图**
4. **选取估计的最大天空亮度值作为迭代阈值，增益大小为 0.1**

a） **迭代计算**

Repeat

并行在各个滤波后的图像中：

(1)找到最亮的点，记录峰值亮度大小，位置及对应的尺度的大小，并计算最亮值与脏束（）卷积再乘以增益（gain）的结果，存储计算结果；

(2)在图像中减去(1）中的计算结果，得到残余图像;

Until

最大亮度值达到阈值 () 或达到预设迭代次数 200

b） **得到洁化图**

在各个尺度中：

（1） 与结束卷积得到洁化图：，，为洁束

（2） 将残余图像加到洁化图中：

将不同尺度得到的不同频率范围的洁图进行叠加，得到最终结果

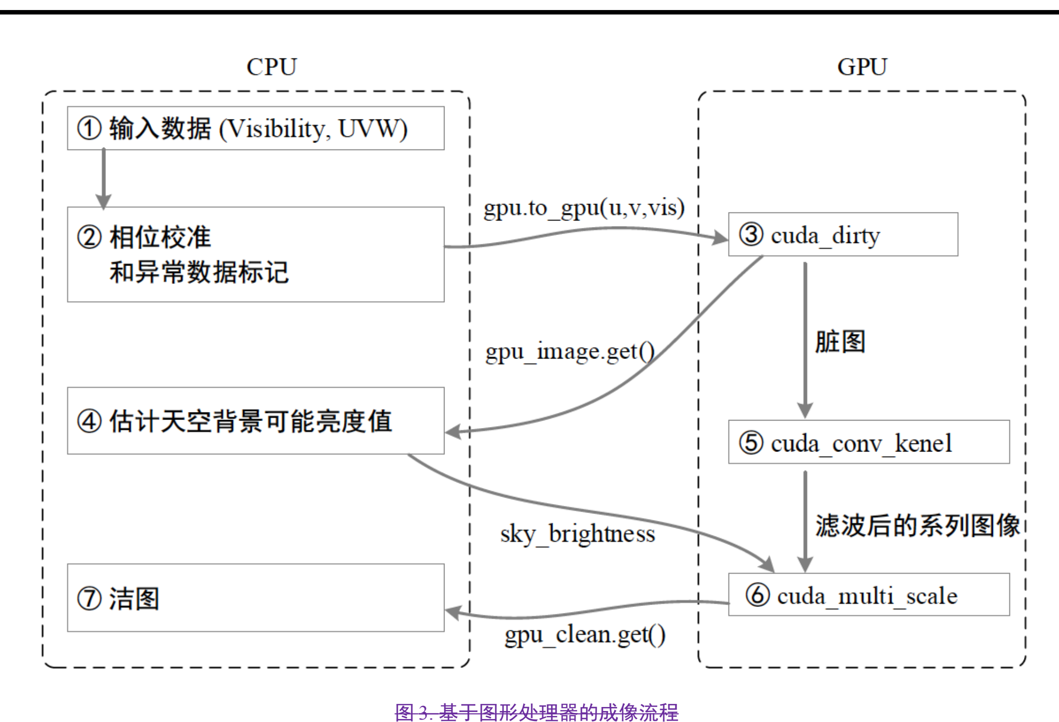
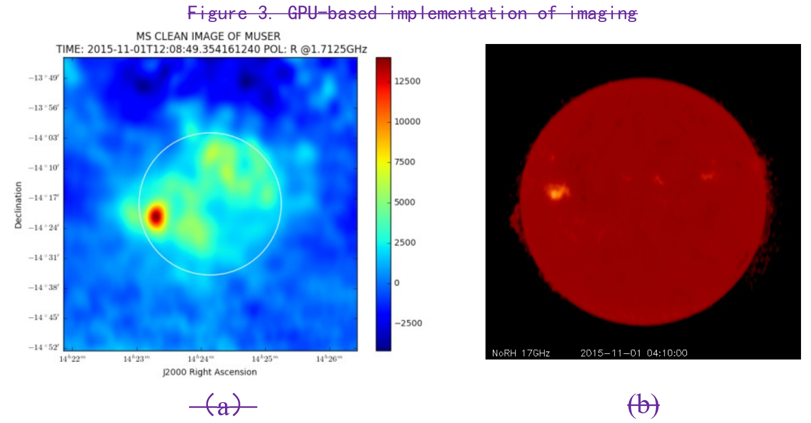


图 3. 基于图形处理器的成像流程

Figure 3. GPU-based implementation of imaging



(a) (b)

图 4(a) MUSER 洁化图(Multi-scale)（MUSER-I at 2015-11-01 04:08:49.354161240 (UTC), 频率: 1.7125 GHz, 右旋）(b) 野边山天文台洁化图（2015-11-01 at 04:10:00 (UTC), frequency:17 GHz (R+L).

Fig.4(a) Clean image of MUSER on 2015-11-01 at 04:08:49.354161240 (UTC), frequency: 1.7125 GHz, polarization: right); (b) Clean image from Nobeyama on 2015-11-01 at 04:10:00 (UTC), frequency:17 GHz (R+L).

3.2 算法效率

由于多个滤波后的图像同时进行洁化, 每个滤波后的图像有不同的迭代次数，且洁化算法的迭代次数动态改变，因此本文仅以具体实例说明多尺度带通滤波的性能。在 NVIDIA Corporation GM200 测试环境下，对于处理图 4 中的 2015-11-011.7125GHz 的一张使用基于带通滤波的多尺度洁化时间为 1.424 s，前期实验中的 Högbom CLEAN 算法[89]的洁化时间为 4.03748 s。经过多次洁化处理实验，表明当前的 Multi-Scale CLEAN 能有效将成图效率提高近 3

，同时可以减少迭代次数，对数据处理流水线性能的提升有重要意义。

4.总结

本文在分析 Multi-Scale CLEAN 算法原理的基础上，实现了基于带通滤波的多尺度洁化算法，较 Cornwell 的多尺度洁化算法在全局不同尺度的图像中的迭代而言，完全实现了不同尺度的并行。实验选取了针对 MUSER 成像的滤波器，尺度大小及不同尺度的取值。同时，基于图形处理器的实现表明，基于带通滤波的多尺度洁化算法比同等条件下的 Högbom CLEAN 性能提高了近三倍，并且可有效减少迭代次数。本文的研究结果将有效提高整个数据处理系统的性能，并对射电干涉成像提供了有价值的参考。

参考文献:

[1] Yan Yihua, Zhang Jian, Chen Zhijun, et al. Progress on Chinese spectral radioheliograph—CSRH construction[C]// Proceedings of the General Assembly and Scientific Symposium. 2011 .

[2] Taylor G B, Carilli C L, Perley R A. Synthesis imaging in radio astronomy II[C]// —ASP Conference Series. 1999.

[3] Cornwell T J. Multiscale CLEAN deconvolution of radio synthesis images [J]. IEEE Journal of Selected Topics in Signal Processing, 2008, 2(5): 793-801.

[4] Koshiishi H. Restoration of solar images by the Steer algorithm [J]. Astronomy & Astrophysics, 2003, 412(3): 893-896.

[5] MUSCAT D. High-Performance Image Synthesis for Radio Interferometry [J]. arXiv preprint arXiv:14034209, 2014.

[56] Mercier C, Subramanian P, Kerdraon A, et al. Combining visibilities from the giant meterwave radio telescope and the Nancay radio heliograph-High dynamic range snapshot images of the solar corona at 327 MHz [J]. Astronomy & Astrophysics, 2006, 447(3): 1189-201.

[67] 卫守林, 刘鹏翔, 王锋, 等. 基于 Spark Streaming 的明安图射电频谱日像仪实时数据处理 [J]. 天文研究与技术, 2017, 14(4): 421-428.
Wei Shoulin, Liu Pengxiang, Wang Feng,et al. Real-time data processing in Mingantu Ultrawide Spectral Radio Heliograph based on Spark Streaming[J]. Astronomical Research & Technology, 2017,14(4): 421-428.

[78] 陈泰燃, 王威, 王锋, 等. 基于 MPI 的高性能 UVFITS 数据合成研究与应用[J]. 天文研究与技术, 2016, 13(2): 184-189.
Chen Tairan, Wang Wei, Wang Feng,et al. The study and application of a high performance UVFITS assembly system based on MPI [J]. Astronomical Research & Technology, 2016,13(2): 184-189.

[89] Wang F, Mei Y, Deng H, et al. Distributed data-processing pipeline for Mingantu Ultrawide Spectral Radioheliograph [J]. Publications of the Astronomical Society of the Pacific, 2015, 127(950): 383-396.

A CLEAN Algorithm Based on Multi-Scale bandpass filter

and its GPU implemention

Ying Mei[1,4], Hui Deng[3], Xiaoli Zhang[2],Feng Wang[1,3,4],Tengda Chen[2]
(1. Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China; 2. Kunming University of Science and Technology, Kunming 650500, China; 3. Center of Astrophysics, Guangzhou University, Guangzhou 510006, China; 3. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: China's new generation Radio telescope---The MingantUSpEctralRadioHeliograph (MUSER) generates high-quality radio images with high temporal, high spatial, and high spectral resolutions in the frequency range of 400 MHz~15 GHz. The MUSER which will provide a new observational window on researches such as solar activities, solar electron accelerations. High-performance and high-quality imaging forms a significantlyimportant aspect of MUSER's massive data processing requirements.Referring to the data processing method of the MEUDON observatory in France, we give a detail analysis of multi-scale CLAEN. Appropriate parameters are given for MUSER CLEAN and the implementation based on GPU is presented. The experimental results show that the multi-scale CLAEN is nearly three times faster than the previously realizedHögbom CLEAN, which will improve the performance of the whole data processing system effectively.

Key words: Multi-Scale CLEAN; Radio interferometric imaging; GPU